

# Matías Ignacio Sánchez Muñoz

Santiago de Chile • Sanchezm.matias@gmail.com • (+56 9) 76109904 • [LinkedIn](#) • [Portafolio](#)

## PERFIL PROFESIONAL

### Data Scientist / ML Engineer | Ingeniero Civil en Computación e Informática

Ingeniero de Datos con especialización en Machine Learning, Computer Vision y Procesamiento de Lenguaje Natural (NLP), con experiencia en diseño, entrenamiento y evaluación de modelos predictivos orientados a producción. Construyo soluciones end-to-end: desde la ingesta y transformación de datos hasta el despliegue de modelos como aplicaciones web. Cuento con el certificado AWS Cloud Practitioner, utilizo infraestructura cloud (AWS, GCP) como plataforma de soporte para pipelines de datos y ejecución de modelos. Combino rigor estadístico curva ROC, matriz de confusión,  $R^2$ , MAE, RMSE con visión de negocio para generar impacto medible en retención de clientes, diagnóstico y automatización inteligente.

## EXPERIENCIA

### ORION 2000 SPA

Mar 2025 – May 2025

Data Analyst / ML Intern

- Desarrollé un modelo predictivo de consumo multicloud (AWS, Azure, GCP) usando Random Forest como algoritmo principal, alcanzando 95 % de precisión con validación cruzada y split 80/20, aplicando feature engineering sobre patrones de gasto histórico de clientes.
- Implementé una prueba de concepto (POC) con interfaz interactiva en Dash, permitiendo filtrar predicciones por tipo de cliente y horizonte temporal, desplegada en Render.
- Construí reportes automatizados en Power BI conectado a Mongo DB con alertas programadas, y ejecuté análisis exploratorios y limpieza de datos con comportamientos anómalos.

### Proyecto de Título – UTEM

Sep 2024 – Ene 2025

Data Scientist / Investigador — Sistema de Inferencia Difusa para Predicción de Churn

- Procesé y transformé un dataset real de 69.403 registros de una plataforma de streaming global, ejecutando feature engineering (frecuencia de uso por mediana y promedio) para alimentar un modelo de inferencia difusa Mamdani con 27 reglas lingüísticas y 3 variables de entrada.
- Medí correlaciones clave: frecuencia de uso → satisfacción ( $r = 0.81$ ), tiempo de suscripción → satisfacción ( $r = 0.60$ ), con métricas MAE = 42.43 y RMSE = 45.00, validando el comportamiento del modelo con análisis de fuzzy strength por segmento.
- Identifiqué 4.475 clientes en riesgo de churn (retorno potencial USD 837.998/mes) y 32.254 candidatos a upgrade (USD 131.093/mes adicionales), entregando una herramienta de BI accionable para campañas de retención y upselling.
- Desplugué el modelo como aplicación web con Django, generando dashboards interactivos (Matplotlib, Chart.js) con visualización personalizada por segmento de cliente.

## PROYECTOS DESTACADOS

### Sistema de Diagnóstico Médico con NLP (Computer Vision & NLP)

2025

- Diseñé e implementé una aplicación web de asistencia de voz para diagnósticos médicos, integrando modelos de Procesamiento de Lenguaje Natural (NLP) para interpretar síntomas descritos en lenguaje natural y sugerir diagnósticos probables.
- Construí el pipeline completo: preprocesamiento de texto, vectorización, clasificación y despliegue como app web en Django con interfaz de voz.

### Seguridad Biométrica — Reconocimiento Facial y Dactilar (CNN)

2025

- Implementé un sistema biométrico multimodal combinando reconocimiento facial mediante redes neuronales convolucionales (CNN) y reconocimiento de huella dactilar, integrando ambos modelos en una aplicación móvil Flutter.
- Apliqué técnicas de Computer Vision para preprocesamiento de imágenes, extracción de características y evaluación del modelo con matriz de confusión y curva ROC.

### Análisis Predictivo de Valor de Propiedades Residenciales

2024

- Ejecuté pipeline completo de ML supervisado y no supervisado sobre datos socioeconómicos y estructurales: análisis exploratorio, reducción dimensional con PCA, clasificación con KNN, SVM, regresión logística y Random Forest, clustering con K-Means y jerárquico (dendrograma).
- Evalué modelos con métricas estándar:  $R^2$ , MAE, RMSE, matriz de confusión y curva ROC, documentando comparativa de algoritmos por rendimiento y complejidad computacional (Big-O).

## Agentes de IA y RAG con LangChain / LangGraph

2025 – 2026

- Arquitecté sistemas multi-agente con RAG (Retrieval-Augmented Generation) integrando bases de datos vectoriales, memoria persistente y modelos de lenguaje (LLMs) para recuperación semántica sobre documentos no estructurados.
- Construí y expuse los flujos como APIs REST con FastAPI, desplegadas con soporte cloud.

## STACK TECNOLÓGICO

### Machine Learning & Data Science

• Python (Pandas, NumPy, Scikit-learn)	• Redes neuronales / Deep Learning (MLP, ANN, rnn)	• Computer Vision (finetuning usando Hugging Face)
• NLP (procesamiento de texto, voz)	• Random Forest / Árboles de decisión	• SVM / KNN / Naive Bayes
• Regresión lineal / logística / Lasso	• K-Means / Clustering jerárquico	• PCA / Reducción dimensional
• Sistemas difusos (Mamdani)	• Series de tiempo	• TensorFlow / Keras

### Evaluación de Modelos

• Matriz de confusión	• Curva ROC / AUC	• R <sup>2</sup> / MAE / RMSE
• Validación cruzada	• Fuzzy Strength / activación	• Análisis de correlación (r)
• Notación Big-O	• Split train/test (80/20)	• Feature engineering

### AWS Cloud (re:Start · Cloud Practitioner) & Infraestructura

• S3 (almacenamiento datasets)	• EC2 / Lambda	• RDS / DynamoDB
• IAM / VPC / CloudWatch	• AWS CLI	• GCP (BigQuery / GCS)
• Docker, Kubernetes K8s	• Linux / Bash scripting	• SSH / conexión remota EC2
• Git / GitHub	• FastAPI	• Terraform / AWS CloudFormation

### Visualización & Despliegue

• Power BI / DAX	• Dash / Streamlit	• Matplotlib / Seaborn
• Chart.js	• Django (web app)	• Excel (análisis de datos)
• SQL (PostgreSQL)	• SQL Server	• MongoDB (NoSQL)

### IA Generativa & Automatización

• LangChain / LangGraph	• RAG (bases vectoriales)	• n8n / MCP
• Agentes multi-agente	• FastAPI (APIs de modelos)	• Render (despliegue)

## EDUCACIÓN & CERTIFICACIONES

### Universidad Tecnológica Metropolitana (UTEM) *carrera de 12 semestres*

marzo 2020 – diciembre 2025

Ingeniero Civil en Computación, Mención en Informática

### AWS Generation – Programa re:Start 480 hrs. 14 semanas

abril 2025 – Julio 2026

AWS Certified Cloud Practitioner | EC2, S3, Lambda, RDS, VPC, IAM, CloudFormation, etc.

### Udemy – Agentes de IA con LangChain y LangGraph

agosto 2025 – enero 2026

Agentes IA, RAG, arquitecturas multi-agente, FastAPI y Streamlit

### Udemy – Automatización con n8n y Model Context Protocol

mayo 2025 – agosto 2025

Orquestación de flujos de datos e integración de LLMs con despliegue cloud

## IDIOMAS

Español (nativo) • inglés A2 – Universidad Central (Jul – Oct 2025, completado)